Niveau(x) d'enseignement: 3 Numéro EDUQ : 8670388
Centre d'intérêt: 540

Brunet, Richard.

Le rôle des tests dans SAGE : une mesure de l'atteinte d'objectifs pédagogiques ou une mesure du rendement aux unités d'apprentissage? / Richard Brunet. -- [Sainte-Foy] : INRS-Education, 1974.
9 f. ; 28 cm.

Titre de la couv. Code: D-22.

Vedette(s)-secondaire(s): INRS-Education

Descripteur(s): Elaboration de tests * Evaluation de tests * Evaluation de projets * Primaire (Ordre d'enseignement)

Identificateur(s) de nom: SAGE (Projet)

Identificateur(s) de sujet: Test critère

Dans le projet SAGE (Système d'apprentissage géré par l'étudiant), il est nécessaire de baser la construction des tests critères sur des objectifs pédagogiques très bien définis, de sorte que ces tests puissent être élaborés indépendamment des unités d'apprentissage. Or, en ce qui a trait à l'ensemble des unités d'apprentissage de quatrième année et à une partie des unités d'apprentissage de cinquième année, cette exigence n'a pas été respectée. Cet accroc à la théorie implique une série de conséquences plus ou moins dramatiques que l'auteur analyse dans ce texte. Il montre que les tests élaborés à l'INRS-Education jusqu'en août 1974 n'ont pas les propriétés des tests critères et que leur élaboration a été contaminée par les unités d'apprentissage.

Type de document: r

EDUQ 8670388

INRS - ÉDUCATION

LE RÔLE DES TESTS DANS SAGE:

UNE MESURE DE L'ATTEINTE D'OBJECTIFS PÉDAGOGIQUES

ou

UNE MESURE DU RENDEMENT AUX UNITÉS D'APPRENTISSAGE?

RICHARD BRUNET

8 août 1974.

D-22

Dans un de ses textes, Jean-Pierre Masson démontre clairement, avec preuves à l'appui, la nécessité de baser la construction des tests critères sur des objectifs pédagogiques très bien définis de sorte que ces tests puissent être élaborés indépendamment des unités d'apprentissage. Or, il semble que, en ce qui a trait à l'ensemble des unités d'apprentissage de 4e année et à une partie des unités d'apprentissage de 5e année, cette exigence n'a pas été respectée pour des raisons justifiables et justifiées.

Cependant, cet accroc à la théorie implique une série de conséquences plus ou moins dramatiques, conséquences que nous essaierons d'analyser dans les pages qui suivent.

EVALUATION DU PROJET SAGE

Il a été prévu dans le plan d'évaluation du projet SAGE entre autres l'évaluation des tests critères et l'évaluation des unités d'apprentissage.

Plus spécifiquement, l'évaluation des tests réfère aux études métriques visant à assurer aux tests les qualités scientifiques de base telles l'homogénéité des item mesurant un même objectif, la fidélité, la validité, la sensibilité des items à l'apprentissage, etc.

L'évaluation des unités d'apprentissage réfère aux études sur la difficulté des voies d'apprentissage, de leur relation avec les objectifs, etc.

Nous allons démontrer que ces études ne peuvent pas être réalisées en entier à cause du fait que les tests élaborés pour mesurer le rendement des étudiants ont été contaminés par les unités d'apprentissage.

LE SCHEME EXPERIMENTAL GENERAL

Le schème expérimental général est le suivant:

Pré-test - Unité d'apprentissage - Post-test

L'élève est soumis à un pré-test qui vise à recueillir des données sur la maîtrise d'un objectif d'apprentissage avant l'apprentissage lui-même; puis l'élève entreprend l'unité d'apprentissage et est soumis à un post-test qui vise à déterminer la maîtrise ou non maîtrise des objectifs poursuivis dans l'unité d'apprentissage.

La comparaison des données du pré et post-test fournit des indications sur la sensibilité des items à l'apprentissage et les données des tests eux-mêmes sont utilisées à des fins d'analyse de test.

Cependant, on n'est sûr ni des tests, ni des unités d'apprentissage ! Par conséquent, on ne saura jamais comment interpréter les résultats obtenus: en effet, ces résultats peuvent être dus à des caractéristiques associées soit au test, soit à l'unité d'apprentissage, soit aux deux.

Pour pouvoir solutionner ce système, il est alors absolument nécessaire de poser au moins un postulat.

POSTULATS NECESSAIRES

Pour pouvoir étudier soit les tests, soit les unités d'apprentissage, on doit donc poser l'un ou l'autre des postulats suivants:

a) les tests sont valides
b) les unités d'apprentissage sont valides.

Note : Bormuth postule la validité des unités d'apprentissage.

Si le postulat (a) est posé, il est donc superflu d'étudier la validité des tests; si le postulat (b) est posé, il est également superflu d'étudier la validité des unités d'apprentissage.

En ce qui concerne les unités d'apprentissage, ni Jean-Pierre, ni moi ne sommes prêts à postuler une telle chose; quant aux tests, il aurait pu être possible de postuler leur validité si les conditions essentielles à la construction des tests-critères avaient été respectées.

LA CONSTRUCTION DES TESTS-CRITERES DANS SAGE

Les tests-critères construits pour mesurer l'apprentissage des élèves ont été élaborés à partir des mêmes objectifs qui ont été utilisés pour l'élaboration des unités d'apprentissage; cependant, dans la plupart des cas, ces tests ont été contaminés par les unités d'apprentissage, soit parce que les items ont été ajustés aux moyens utilisés pour faire effectuer l'apprentissage à l'élève, soit parce que des items ont été éliminés à cause d'une interprétation différente d'un même objectif de départ de la part du constructeur de test et du rédacteur-éditeur, laissant supposer une ambiguité de cet objectif, soit parce qu'il n'y avait pas d'objectifs clairement définis

On peut donc en conclure que ces tests ont perdu leurs qualités de tests-critères et il est pour le moins difficile d'en postuler la validité comme on aurait pu le faire s'il s'était agi de véritables tests-critères.

UN EMBRYON DE THEORIE "HOME MADE"

Je vais essayer d'insérer les unités d'apprentissage et les tests dans la théorie générale du "domain sampling". Quoique développée pour les tests normatifs, je crois que cette théorie s'applique

au cas qui nous préoccupe, même si la façon statistique traditionnelle d'évaluer les paramètres ne s'appliquent pas.

A.- LES TESTS

Au point de départ, nous avons un objectif d'apprentissage; le test doit mesurer jusqu'à quel point l'élève maîtrise l'objectif. La façon la plus sûre de déterminer le niveau de maîtrise de l'élève est de lui administrer l'univers des items mesurant la maîtrise de cet objectif. Ce test serait évidemment de longueur infinie, ce qui n'est pas très pratique.

Il est donc nécessaire d'échantillonner un certain nombre d'items de cet univers en essayant d'être le plus représentatif possible de sorte que, au moyen de l'échantillon, on puisse estimer le plus fidèlement possible le score qu'aurait eu l'élève si on lui avait administré l'univers des items.

Si l'objectif de départ est très spécifique (unidimentionnel), il est relativement facile d'échantillonner des items représentatifs. Si par contre cet objectif est vague, deux constructeurs échantillonneront des items en fonction de sous-objectifs plus unidimensionnels de sorte qu'il est fort possible que les deux tests résultant ne soient pas parallèles.

Il est donc essentiel pour assurer une validité intrinsèque aux tests que le domaine échantillonné (objectif) soit le plus spécifique (unidimensionnel) possible.

B.- LES UNITES D'APPRENTISSAGE

Au point de départ, nous avons toujours le même objectif d'apprentissage. La façon la plus sûre d'amener l'étudiant à maîtriser l'objectif est de lui présenter toutes les voies possibles d'apprentissage,

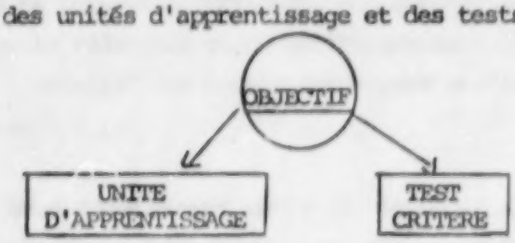
ce qui pratiquement est impensable. Le rédacteur-éditeur doit donc nécessairement choisir dans le domaine des voies possibles, une voie particulière. Si l'objectif est parfaitement précis (unidimensionnel) deux rédacteurs-éditeurs élaboreront des voies parallèles alors que si cet objectif est flou, il est très plausible que les deux voies soient élaborées en fonction de sous-objectifs spécifiques à chaque rédacteur-éditeur, ces voies n'étant alors pas parallèles.

Donc, plus l'objectif de départ sera spécifique, plus on sera sûr que la voie d'apprentissage mène l'élève à une maîtrise réelle de cet objectif.

Il est donc essentiel d'avoir au point de départ les objectifs les plus spécifiques possible.

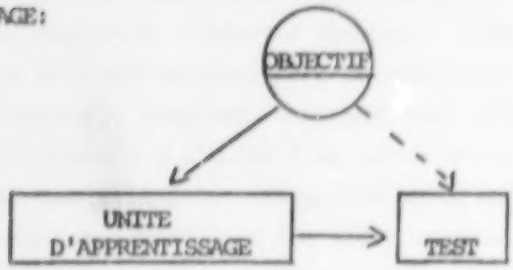
LA CONTAMINATION DES TESTS PAR LES UNITES D'APPRENTISSAGE

La figure ci-dessous représente la structure idéale d'élaboration des unités d'apprentissage et des tests-critères.



Ainsi, à partir d'un objectif donné, le constructeur de test élabore un test-critère indépendamment du rédacteur-éditeur qui rédige l'unité d'apprentissage. Cette indépendance est une condition absolument nécessaire, comme Jean-Pierre le montre dans son texte.

La figure qui suit représente ce qui s'est passé de fait dans SAGE:



Un objectif de départ donne lieu à la création d'une unité d'apprentissage. Cependant, comme l'objectif est plus ou moins précis, le test doit être construit à la fois à partir de l'unité d'apprentissage et de l'objectif, et cela dans des proportions très variables.

Dans l'unité d'apprentissage, l'objectif est opérationnalisé en activités, considérées comme des moyens d'apprentissage. La contamination du test se fait en restreignant son contenu à des sous-objectifs implicites pour le rédacteur et/ou en introduisant dans la mesure de l'atteinte de l'objectif une mesure des moyens utilisés dans l'unité d'apprentissage.

Nous avons donc d'une part une unité d'apprentissage qui comprend une composante "objectif", une composante "moyen" et une composante "erreur", et d'autre part, un test qui comprend lui aussi une composante "objectif", une composante "moyen" et une composante "erreur". Il est donc, impossible d'isoler la composante "objectif" de la composante "moyen" ; concrètement, ceci implique l'impossibilité d'étudier la validité de l'unité d'apprentissage, i.e. sa relation avec l'objectif, puisque cette relation sera toujours surévaluée dans une proportion qui variera d'une unité d'apprentissage à l'autre, de façon directement proportionnelle au degré de contamination. Vrai

Je vais maintenant prouver ce point à l'aide d'un cas extrême. Supposons l'objectif suivant: "donner à l'élève les habiletés mentales nécessaires à la résolution de problèmes de figures renversées". Supposons également une unité d'apprentissage dans laquelle on montrerait à l'élève à se servir d'une surface réfléchissante pour résoudre les problèmes. Supposons enfin un test dans lequel il est possible pour l'élève d'utiliser une telle surface réfléchissante pour résoudre les item

L'administration du test montrerait que l'élève a atteint l'objectif visé alors que l'objectif n'a pas du tout été atteint. Cette conclusion vient du fait que la composante commune au test et à l'unité d'apprentissage est le "moyen" plutôt que la "fin ".

On pourra toujours servir l'argument que le rédacteur-éditeur et le constructeur de test ne sont pas idiots (ce que j'admets sans preuve expérimentale...je le postule), et que ceux-ci ne perdront pas complètement de vue la fin pour ne s'attarder qu'au moyen. Il existe toutefois amplement de littérature qui démontre l'existence des biais systématiques ainsi que leur subtilité, et je me refuse à croire à l'absence totale de tels biais dans SAGE.

Si ce bais était constant d'un test à l'autre, il aurait été possible de solutionner partiellement le problème en utilisant une norme quelconque; comme son importance varie d'un test à l'autre, il devient absolument impossible de pouvoir interpréter les résultats.

TEST CRITERIE OU TEST "UNITE MINIATURE"

Placé devant le fait où il devient impossible de traiter les tests déjà développés comme des tests critériés puisque ceux-ci en ont perdu les propriétés, nous nous trouvons devant une situation déplaisante: les tests ne pourront servir qu'à évaluer la performance de

l'élève à une voie d'apprentissage particulière et non pas à un objectif d'apprentissage dans son sens strict. Ou bien on définit l'objectif par la voie d'apprentissage (la fin par les moyens), ou bien on construit des tests critériés tels qu'ils devraient être construits. Comme ni l'un ni l'autre n'est réalisable à court terme (le premier impliquant soit une redéfinition des objectifs et une re-rédaction des voies, soit le statu quo qui n'arrange rien) il sera impossible d'étudier la "validité" des unités d'apprentissage en fonction de l'objectif dans la mesure où il y a contamination du test par la voie d'apprentissage.

De l'avis de Jean-Pierre, les tests pour les 32 sous-modules de français (5^{e} année) ont subi très peu de contamination alors que les 41 modules de français ont été partiellement contaminés et que ceux des 54 modules (4^{e} année) l'ont été fortement.

Dans la mesure où les objectifs d'apprentissage peuvent porter à interprétation, il est impossible de construire des tests qui ont les propriétés des test critères. On doit alors construire des tests "unité miniature" consistant à reproduire en miniature les activités d'apprentissage contenues dans l'unité. Si l'élève peut reproduire fidèlement les tâches on lui accorde la note de passage, sinon l'élève doit accomplir d'autres tâches qu'on postule reliées aux objectifs.

Vrai

EN RESUME

Certains accrocs théoriques font que, pour la majorité des unités d'apprentissage, les études prévues ne pourront être effectuées et que les tests élaborés jusqu'à ce jour n'ont pas les propriétés des tests critères. Je souhaite vivement pouvoir discuter de ce problème avec les gens intéressés.